

A GENERALISATION OF TANAKA'S FORMULA

GIANLUCA CASSESE

ABSTRACT. We prove some new results related to Tanaka's formula.

1. INTRODUCTION.

Since the early treatment by Meyer in [5] and later in greater detail by Dellacherie and Meyer [3], the classical formulas of Itô and Tanaka are among the cornerstones of stochastic analysis. Roughly put, these state that semimartingales are a stable class with respect to twice continuously differentiable and to convex transformations respectively and provide a detailed description of the characteristics of the transformed semimartingale in terms of those of the original one.

In this paper we obtain some new results in this direction as well as some new proofs of known facts. Our starting point is a class of transformations of a given semimartingale X , the X -summable functions, which are our model for treating the evolution of quantities along with X . Included in this class are the functional transformations of X , $f(X)$, as well as the stochastic integrals driven by X . The key technical step is a general representation theorem for linear functionals based on the notion of conglomerability and developed in [2]. This permits to obtain, in Theorem 1, an explicit representation of the expected value of X -summable functions at points in time. From this general formula we then deduce more explicit implications, such as Theorem 4 proving that a \mathcal{C}^1 transformation f of a semimartingale is again a semimartingale and providing its explicit expansion for the case in which the derivative of f is locally Lipschitz.

The novelty of our approach is that, although we aim at final applications which are entirely expressed in the language of the classical theory of probability, we choose to exploit the full-fledged generality of the finitely additive integral in order to overcome some technical difficulties with measurability that would otherwise represent quite an issue. This somewhat unorthodox strategy is loosely reminiscent of the one often followed in problems involving polynomial equations, in which it is often useful to work with fields which are extensions of the one in which a solution is sought for.

Throughout the paper (Ω, \mathcal{F}, P) will be a classical probability space endowed with a filtration $(\mathcal{F}_t : t \in \mathbb{R}_+)$ satisfying the usual assumptions and X a right continuous semimartingale. $\langle X^c, X^c \rangle$ will be the angle bracket of the continuous, local martingale part of X and, in the special case in which $[X, X]$ is locally integrable, $\langle X, X \rangle$ will be the predictable compensator of $[X, X]$. m_X will

Date: April 3, 2017.

2010 Mathematics Subject Classification. Primary: 60G07, 60H05; Secondary: 60J55.

Key words and phrases. Conglomerability, Itô's lemma, Semimartingale, Stochastic calculus, Tanaka's formula.

indicate the countably additive set function induced by $\langle X, X \rangle$ on $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)$ and defined as

$$(1) \quad m_X(H) = P \int \mathbb{1}_H d\langle X, X \rangle \quad H \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+).$$

Eventually, \mathcal{T} will be the family of stopping times τ satisfying $P(\tau < \infty) = 1$ and X^τ the process X stopped at $\tau \in \mathcal{T}$. A family $\{T(n, k) : n, k \in \mathbb{N}\}$ in \mathcal{T} such that $T(0, k) = 0$, $T(n-1, k) < T(n, k)$ on $\{T(n, k) > 0\}$ and $P(\sup_n T(n, k) < \infty) = 1$ for all $k \in \mathbb{N}$ and $\lim_k \sup_n [T(n, k) \wedge t - T(n-1, k) \wedge t] = 0$ for all $t \in \mathbb{R}_+$ is a Riemann sequence. \mathcal{P}_0 will be the collection of predictable rectangles and \mathcal{P} and $\sigma\mathcal{P}$ the ring and σ ring generated by \mathcal{P}_0 , respectively.

We write the expected value of a measurable quantity f somehow unconventionally, as $P(f)$. We denote with the symbols $\mathfrak{F}(A)$ and $\mathfrak{B}(A)$ the class of real valued functions on some set A and the subset of all bounded functions, respectively.

2. MAIN REPRESENTATION

Our starting point is the class \mathcal{H}_X^0 of X -summable functions, i.e. those $h \in \mathfrak{F}(\Omega \times \mathbb{R} \times \mathbb{R})$ such that: (i) $h(\omega, x, x) = 0$ for all $(\omega, x) \in \Omega \times \mathbb{R}$ and (ii) for all $\tau \in \mathcal{T}$ the limit

$$(2) \quad I_X(h)_\tau = \lim_k \sum_n h(X_{T(n-1,k)}^\tau, X_{T(n,k)}^\tau)$$

exists in probability and is independent of the intervening Riemann sequence $\langle T(n, k) \rangle_{k \in \mathbb{N}}$. In general, when treating $h \in \mathcal{H}_X^0$ we will omit reference to Ω but it is important to notice that $h \in \mathcal{H}_X^0$ and $b \in \mathfrak{F}(\Omega)$ imply $hb \in \mathcal{H}_X^0$ with $I_X(hb) = I_X(h)b$. When $h \in \mathcal{H}_X^0$ we define

$$(3) \quad \mathcal{T}(h) = \{\tau \in \mathcal{T} : I_X(h)_\tau \in L^1(P)\}$$

and construct the families $\mathcal{P}_0(h)$ and $\mathcal{P}(h)$ similarly to \mathcal{P}_0 and \mathcal{P} with $\mathcal{T}(h)$ in place of \mathcal{T} . It is useful to remark that $\mathcal{T}(h)$ is a lattice and $\mathcal{P}_0(h)$ a lattice of sets.

An element of \mathcal{H}_X^0 of special importance is $h_0(x, y) = (y - x)^2$. Obviously, $I_X(h_0)_t = [X, X]_t$. In general, for each $h \in \mathcal{H}_X^0$, the quantity $I_X(h)$ describes a (not necessarily adapted) process starting at $I_X(h)_0 = 0$ and describing the increments of h along X . Heuristically, it is tempting to interpret $I_X(h)$ as a sort of generalized stochastic integral with respect to X . Easy examples of the sums in (2) are (a) $h(x, y) = f(y) - f(x)$ or (b) $h(x, y) = g(x)(y - x)$. In the former case, we simply get $I_X(h)_\tau = f(X_\tau) - f(X_0)$; in the latter we have $I_X(h)_\tau = \int_0^\tau (g \circ X)_- dX$ when g is locally bounded and measurable.

In order to introduce doubly indexed processes, we let

$$(4) \quad \tilde{\Omega}_n = \{(\omega, t, s) \in \Omega \times \mathbb{R}_+^2 : 0 < t - s < 1/n\} \quad \text{and} \quad \tilde{\Omega} = \tilde{\Omega}_0.$$

Setting conventionally $0/0 = 0$, we define the positive, linear map $T_X : \mathfrak{F}(\Omega \times \mathbb{R}^2) \rightarrow \mathfrak{F}(\tilde{\Omega})$ via

$$(5) \quad T_X(f)(\omega, t, s) = \frac{f(\omega, X_s(\omega), X_t(\omega))}{(X_s(\omega) - X_t(\omega))^2} \quad (\omega, t, s) \in \tilde{\Omega}.$$

In order to have a nicer mathematical structure than \mathcal{H}_X^0 we define

$$(6) \quad \mathcal{H}_X = \{h \in \mathcal{H}_X^0 : |h| \leq h' \text{ for some } h' \in \mathcal{H}_X^0\},$$

the associated vector space¹

$$(7) \quad \mathcal{V}_X = \left\{ \sum_{i=1}^I T_X(h_i) \mathbf{1}_{] \sigma_i, \tau_i] \times \mathbb{R}_+} : h_i \in \mathcal{H}_X, \sigma_i, \tau_i \in \mathcal{T}(h_i) \ i = 1, \dots, I, \ I \in \mathbb{N} \right\}$$

and the generated ideal

$$(8) \quad \mathcal{L}_X = \{g \in \mathfrak{F}(\tilde{\Omega}) : |g| \leq V \text{ for some } V \in \mathcal{V}_X\}.$$

Notice that $\mathcal{V}_X \subset \mathcal{L}_X$ because of the way we defined \mathcal{H}_X . In the applications, proving the inclusion $h \in \mathcal{H}_X$ will be a major step.

Theorem 1. *Let X be locally bounded. There exist (i) a positive linear functional ψ on \mathcal{L}_X and (ii) a positive, linear map $(\cdot)^{\mathcal{P}} : \mathcal{L}_X \rightarrow L^1(\Omega \times \mathbb{R}_+, \sigma\mathcal{P}, m_X)$ which satisfy the conditions*

$$(9a) \quad \lim_n \inf_{(\omega, t, s) \in \tilde{\Omega}_n} g(\omega, t, s) > -\infty \text{ implies } \psi(g) \geq 0 \text{ and}$$

$$(9b) \quad \lim_n \inf_{(\omega, t, s) \in \tilde{\Omega}_n} (g'(\omega, t, s) - g(\omega, t, s)) \geq 0 \text{ implies } (g \mathbf{1}_B)^{\mathcal{P}} \leq (g')^{\mathcal{P}} \mathbf{1}_B \text{ } m_X \text{ a.s., } B \in \sigma\mathcal{P}$$

and such that the following representation holds:

$$(10) \quad P(I_X(h)_\tau) = \psi(T_X(h) \mathbf{1}_{]0, \tau] \times \mathbb{R}_+}) + P \int_0^\tau T_X^{\mathcal{P}}(h) d\langle X, X \rangle \quad \tau \in \mathcal{T}(h), \ h \in \mathcal{H}_X.$$

In (10), $T_X^{\mathcal{P}}(h)$ is a predictable process associated with $h \in \mathcal{H}_X$ and such that

$$(11) \quad T_X^{\mathcal{P}}(h) \mathbf{1}_{]0, \tau]} = (T_X(h) \mathbf{1}_{]0, \tau] \times \mathbb{R}_+})^{\mathcal{P}} \quad \tau \in \mathcal{T}(h).$$

Proof. Of course, if $[X, X]_\infty = 0$ a.s. then $I_X(h) = T_X(h) = 0$ and (10) becomes trivial. We shall therefore assume throughout that $P([X, X]_\infty > 0) > 0$.

$[X, X]$ is locally integrable. Assume first that $[X, X]_\infty \in L^1(P)$: thus, $\mathcal{P}(h_0) = \mathcal{P}$. Fix $h \in \mathcal{H}_X$. The decomposition property

$$h(X_s^{\sigma \wedge \tau}, X_t^{\sigma \wedge \tau}) + h(X_s^{\sigma \vee \tau}, X_t^{\sigma \vee \tau}) = h(X_s^\sigma, X_t^\sigma) + h(X_s^\tau, X_t^\tau) \quad \sigma, \tau \in \mathcal{T}$$

extends to $P(I_X(h)_\tau)$ so that the following writing

$$(12) \quad \phi(h;]0, \tau]) = P(I_X(h)_\tau) \quad \tau \in \mathcal{T}(h)$$

implicitly defines a strongly additive set function on the lattice $\mathcal{P}_0(h)$ which may be extended to an additive set function to $\mathcal{P}(h)$, [1, theorem 3.1.6], by letting

$$(13) \quad \phi(h;]\sigma, \tau]) = \phi(h;]0, \tau]) - \phi(h;]0, \sigma]) \quad \sigma, \tau \in \mathcal{T}(h).$$

¹ More precisely, by writing $T_X(h) \mathbf{1}_{] \sigma, \tau] \times \mathbb{R}_+}$ we mean the element of $\mathfrak{F}(\tilde{\Omega})$ such that $(T_X(h) \mathbf{1}_{] \sigma, \tau] \times \mathbb{R}_+})(\omega, t, s) = T_X(h)(\omega, t, s) \mathbf{1}_{] \sigma, \tau]}(\omega, t)$.

Fix $I \in \mathbb{N}$ and for $i = 1, \dots, I$ let $h_i \in \mathcal{H}_X$ and $\sigma_i, \tau_i \in \mathcal{T}(h_i)$. We claim that

$$(14) \quad \sum_i \phi(h_i;]\sigma_i, \tau_i]) < 0 \quad \text{implies} \quad \lim_k \inf_{(\omega, t, s) \in \tilde{\Omega}_k} \sum_i T_X(h_i) \mathbb{1}_{]\sigma_i, \tau_i] \times \mathbb{R}_+} < 0.$$

In fact, assume that $\sum_{i=1}^I \phi(h_i;]\sigma_i, \tau_i]) \leq -3a < 0$. According to (2), along any Riemann sequence $\langle \tau_{n,k} \rangle_{k \in \mathbb{N}}$ and for all k sufficiently large, the inequality

$$(15) \quad -2a > \sum_i \sum_n h_i(X_{\tau_{n-1,k}}^{\tau_i}, X_{\tau_{n,k}}^{\tau_i}) - h_i(X_{\tau_{n-1,k}}^{\sigma_i}, X_{\tau_{n,k}}^{\sigma_i})$$

obtains with positive probability. Choosing the Riemann sequence so as to include all times σ_i and τ_i for $i = 1, \dots, I$, (15) simplifies into

$$(16) \quad \begin{aligned} -2a &> \sum_n \sum_{\{i: \sigma_i < \tau_{n,k} \leq \tau_i\}} h_i(X_{\tau_{n-1,k}}, X_{\tau_{n,k}}) \\ &= \sum_n (X_{\tau_{n-1,k}} - X_{\tau_{n,k}})^2 \sum_i \frac{h_i(X_{\tau_{n-1,k}}, X_{\tau_{n,k}})}{(X_{\tau_{n-1,k}} - X_{\tau_{n,k}})^2} \mathbb{1}_{\{\sigma_i < \tau_{n,k} \leq \tau_i\}} \\ &= \sum_n (X_{\tau_{n-1,k}} - X_{\tau_{n,k}})^2 \sum_i T_X(h_i)(\tau_{n,k}, \tau_{n-1,k}) \mathbb{1}_{\{\sigma_i < \tau_{n,k} \leq \tau_i\}} \\ &\geq \sum_n (X_{\tau_{n-1,k}} - X_{\tau_{n,k}})^2 \inf_{0 < t-s < 1/k} \sum_i (T_X(h_i) \mathbb{1}_{]\sigma_i, \tau_i] \times \mathbb{R}_+})(t, s). \end{aligned}$$

Taking limits in probability as $k \rightarrow \infty$, the following inequality must hold with positive probability:

$$-a > [X, X]_\infty \lim_k \inf_{0 < t-s < 1/k} \sum_i T_X(h_i) \mathbb{1}_{]\sigma_i, \tau_i] \times \mathbb{R}_+},$$

which proves our preceding claim.

Let $g \in \mathfrak{F}(\tilde{\Omega})$ and $\sum_i T_X(h_i) \mathbb{1}_{]\sigma_i, \tau_i] \times \mathbb{R}_+} \in \mathcal{V}_X$ satisfy

$$(17) \quad g \sim \sum_i T_X(h_i) \mathbb{1}_{]\sigma_i, \tau_i] \times \mathbb{R}_+} \quad \text{i.e.} \quad \lim_k \sup_{(\omega, t, s) \in \tilde{\Omega}_k} \left| g - \sum_i T_X(h_i) \mathbb{1}_{]\sigma_i, \tau_i] \times \mathbb{R}_+} \right| = 0.$$

Then, using (14), we conclude that writing

$$(18) \quad F(g) = \sum_i \phi(h_i;]\sigma_i, \tau_i]) \quad g \sim \sum_i T_X(h_i) \mathbb{1}_{]\sigma_i, \tau_i] \times \mathbb{R}_+}$$

implicitly defines a positive linear functional on the vector space

$$(19) \quad \tilde{\mathcal{V}}_X = \{g \in \mathfrak{F}(\tilde{\Omega}) : g \sim V \text{ for some } V \in \mathcal{V}_X\}.$$

Let $\tilde{\mathcal{L}}_X$ be the induced ideal, so that $\mathcal{V}_X \subset \tilde{\mathcal{V}}_X$ and $\mathcal{L}_X \subset \tilde{\mathcal{L}}_X$.

Given that $\tilde{\mathcal{V}}_X$ is an ordered vector space containing the constants and that, in the terminology introduced in [2], F is conglomerative with respect to the identity on $\tilde{\mathcal{V}}_X$, we conclude from [2, theorem 1] that there exist: (i) a positive, additive set function μ defined on all subsets of $\tilde{\Omega}$, satisfying $\tilde{\mathcal{L}}_X \subset L^1(\mu)$ and (ii) a positive, linear functional ψ on $\tilde{\mathcal{L}}_X$ vanishing on $\mathfrak{B}(\Omega \times \mathbb{R}_+)$ such that

$$(20) \quad F(g) = \psi(g) + \int g d\mu \quad g \in \tilde{\mathcal{V}}_X$$

Given that $g\mathbf{1}_{\tilde{\Omega} \setminus \tilde{\Omega}_k} \sim \mathbf{1}_{\tilde{\Omega} \setminus \tilde{\Omega}_k} \sim 0$ when $g \in \mathcal{L}_X$, we conclude that

$$(21) \quad \psi(g\mathbf{1}_{\tilde{\Omega} \setminus \tilde{\Omega}_k}) = \mu(\tilde{\Omega} \setminus \tilde{\Omega}_k) = 0 \quad k \in \mathbb{N}$$

and

$$(22) \quad \phi(h;]]\sigma, \tau][) = \psi(T_X(h)\mathbf{1}_{]]\sigma, \tau][\times \mathbb{R}_+}) + \int T_X(h)\mathbf{1}_{]]\sigma, \tau][\times \mathbb{R}_+} d\mu \quad h \in \mathcal{H}_X, \sigma, \tau \in \mathcal{T}(h).$$

Clearly, (21) implies (9a); likewise

$$(23) \quad \lim_n \inf_{\omega, t, s \in \tilde{\Omega}_n} (g'(\omega, t, s) - g(\omega, t, s)) \geq 0 \quad \text{implies} \quad \int (g' - g) d\mu \geq 0.$$

Let λ be the \mathcal{P} marginal of μ . Then,

$$\lambda(B) = \mu(B \times \mathbb{R}_+) = \int T_X(h_0)\mathbf{1}_{B \times \mathbb{R}_+} d\mu = \phi(h_0; B) = m_X(B) \quad B \in \mathcal{P}.$$

Therefore $\bar{\lambda} = m_X|_{\sigma\mathcal{P}}$ is the unique, countably additive extension of λ to the generated σ ring. More generally, fix $g \in \mathcal{L}_X$, $g \geq 0$ and define

$$(24) \quad \lambda_g(B) = \int g\mathbf{1}_{B \times \mathbb{R}_+} d\mu \quad B \in \mathcal{P}.$$

Choosing $\varepsilon > 0$ arbitrary and k large enough so that $\int (g \wedge k) d\mu \geq \int g d\mu - \varepsilon$, we conclude

$$\lambda_g(B) = \int g\mathbf{1}_{B \times \mathbb{R}_+} d\mu \leq \varepsilon + k\lambda(B)$$

i.e. $\lambda_g \ll \lambda$. Then also $\bar{\lambda}_g \ll \bar{\lambda}$, with $\bar{\lambda}_g$ the countably extension of λ_g to $\sigma\mathcal{P}$, and we can thus define $(g)^\mathcal{P} \in L^1(\bar{\lambda})$ to be the corresponding Radon Nikodym derivative. (9b) follows from (23) and uniqueness of the Radon Nikodym derivative.

If $h \in \mathcal{H}_X$, let $\langle \tau_n^h \rangle_{n \in \mathbb{N}}$ be an increasing sequence in $\mathcal{T}(h)$ such that

$$\lim_n \mu(]]0, \tau_n^h][\times \mathbb{R}_+) = \sup_{\tau \in \mathcal{T}(h)} \mu(]]0, \tau][\times \mathbb{R}_+)$$

and define

$$(25) \quad T_X^\mathcal{P}(h) = \sum_n (T_X(h)\mathbf{1}_{]]\tau_{n-1}^h, \tau_n^h][\times \mathbb{R}_+})^\mathcal{P}.$$

If $\tau \in \mathcal{T}(h)$ and $B \in \sigma\mathcal{P}$ we deduce (11) from

$$(26) \quad \begin{aligned} \int (T_X(h)\mathbf{1}_{]]0, \tau][\times \mathbb{R}_+})^\mathcal{P} \mathbf{1}_B dm_X &= \lim_n \int_0^\tau (T_X(h)\mathbf{1}_{]]0, \tau_n^h][\times \mathbb{R}_+})^\mathcal{P} \mathbf{1}_B dm_X \\ &= \int_0^\tau \sum_n (T_X(h)\mathbf{1}_{]]\tau_{n-1}^h, \tau_n^h][\times \mathbb{R}_+})^\mathcal{P} \mathbf{1}_B dm_X \\ &= \int T_X^\mathcal{P}(h)\mathbf{1}_{]]0, \tau][} \mathbf{1}_B dm_X. \end{aligned}$$

We obtain (10) by combining (22) with (26).

In the general case $[X, X]$ is integrable only along some localizing sequence $\langle \sigma_j \rangle_{j \in \mathbb{N}}$. We can then establish the existence of a mapping, $T_X^\mathcal{P}(h; j)$ which satisfies (10) in restriction to the interval $]]\sigma_{j-1}, \sigma_j][$. But then it is enough to define $T_X^\mathcal{P}(h) = \sum_j T_X^\mathcal{P}(h; j)\mathbf{1}_{]]\sigma_{j-1}, \sigma_j][}$. \square

It is a novel feature of our approach to work with a doubly indexed stochastic process, such as $T_X(h)$, and under rather weak measurability conditions. These unusual aspects are easily approached via the finitely additive integral representation adopted which has however the drawback that the operator $T_X^{\mathcal{P}}(h)$, despite a superficial resemblance with the classical predictable projection, lacks in fact some desirable properties such as monotone continuity. Moreover the representation (10) contains the term ψ which is difficult to treat explicitly and that we set out to dispose of in the next results.

Theorem 2. *Let $h \in \mathcal{H}_X$ admit a sequence $\langle \sigma_n \rangle_{n \in \mathbb{N}}$ in \mathcal{T} which increases to ∞ a.s. and such that*

$$(27) \quad \lim_{c \rightarrow \infty} P^* \left(\sup_{s < t \leq \sigma_n} |T_X(h)| > c \right) = 0 \quad n \in \mathbb{N}.$$

Then,

$$(28) \quad P(I_X(h)_\tau) = P \int_0^\tau T_X^{\mathcal{P}}(h) d\langle X, X \rangle \quad \tau \in \mathcal{T}(h).$$

In addition, if $I_X(h)$ is adapted and locally integrable it is then of locally integrable variation, with predictable compensator $\int T_X^{\mathcal{P}}(h) d\langle X, X \rangle$. In this special case,

$$(29) \quad I_X(h)_t = \int_0^t T_X^{\mathcal{P}}(h) d\langle X^c, X^c \rangle + \sum_{s \leq t} \Delta I_X(h)_s \quad t \in \mathbb{R}_+.$$

Proof. We first prove the claim in the case in which X is locally bounded. Assume then, with no loss of generality, that $\sigma_n \in \mathcal{T}(h_0)$. Let $\langle A_{n,k} \rangle_{k \in \mathbb{N}}$ be an increasing sequence in \mathcal{F} with $A_{n,k} \subset \{\sup_{s < t \leq \sigma_n} |T_X(h)| \leq k\}$ and $\lim_k P(A_{n,k}^c) = 0$. According to (22), if $\tau \in \mathcal{T}(h)$

$$P(I_X(h)_{\tau \wedge \sigma_n}) = \lim_k P(I_X(h \mathbf{1}_{A_{n,k}})_{\tau \wedge \sigma_n}) = \lim_k \int T_X(h) \mathbf{1}_{A_{n,k}} \mathbf{1}_{[0, \tau \wedge \sigma_n]} \times \mathbb{R}_+ d\mu$$

i.e.

$$(30) \quad P(I_X(h)_{\tau \wedge \sigma_n}) = \int T_X(h) \mathbf{1}_{[0, \tau \wedge \sigma_n]} \times \mathbb{R}_+ d\mu = P \int_0^{\tau \wedge \sigma_n} T_X^{\mathcal{P}}(h) d\langle X, X \rangle$$

because

$$\lim_k \int \mathbf{1}_{A_{n,k}^c} \mathbf{1}_{[0, \tau \wedge \sigma_n]} \times \mathbb{R}_+ d\mu = \lim_k \int \mathbf{1}_{A_{n,k}^c} \mathbf{1}_{[0, \tau \wedge \sigma_n]} \times \mathbb{R}_+ T_X(h_0) d\mu = \lim_k P(I_X(h_0)_{\tau \wedge \sigma_n} \mathbf{1}_{A_{n,k}^c}) = 0.$$

In addition, we deduce from (30) the inequality

$$(31) \quad \left| P(I_X(h)_{\tau \wedge \sigma_j}) - P(I_X(h)_{\tau \wedge \sigma_{j+p}}) \right| \leq \int |T_X(h)| \mathbf{1}_{[\tau \wedge \sigma_j, \tau \wedge \sigma_{j+p}]} \times \mathbb{R}_+ d\mu$$

and from this in turn

$$(32) \quad P(I_X(h)_\tau) = \int T_X(h) \mathbf{1}_{[0, \tau]} \times \mathbb{R}_+ d\mu = P \int_0^\tau T_X^{\mathcal{P}}(h) d\langle X, X \rangle.$$

Fix $\sigma, \tau \in \mathcal{T}$ and let $b \in \mathfrak{B}(\mathcal{F})$ be such that $|I_X(h)_\tau - I_X(h)_\sigma| = I_X(hb)_\tau - I_X(hb)_\sigma$. Then,

$$P(|I_X(h)_\tau - I_X(h)_\sigma|) = P \int_\sigma^\tau T_X^{\mathcal{P}}(hb) d\langle X, X \rangle \leq P \int_\sigma^\tau |T_X^{\mathcal{P}}(h)| d\langle X, X \rangle$$

so that indeed $I_X(h)$ is of locally integrable variation. If $I_X(h)$ is adapted, then from classical results, [3, VI.13], we further deduce that $I_X(h)^\tau - \int T_X^\mathcal{P}(h)d\langle X^\tau, X^\tau \rangle$ is a uniformly integrable martingale for each $\tau \in \mathcal{T}$. We can then write

$$(33) \quad I_X(h)_\tau = M_\tau + \int_0^\tau T_X^\mathcal{P}(h)d\langle X, X \rangle \quad \tau \in \mathcal{T}$$

with M a pure jump martingale. Thus, $\sum_{s \leq t} \Delta I_X(h)_s = M_t + \sum_{s \leq t} T_X^\mathcal{P}(h)_s \Delta X_s^2$ from which (29) readily follows. This proves the claim for the case in which X is locally bounded.

Returning to the general case, fix $a > 0$ and define $X_t^a = X_t - \sum_{s \leq t} \Delta X_s \mathbf{1}_{\{|\Delta X_s| > a\}}$. Clearly X^a satisfies (29), since locally bounded. Moreover, for each $T > 0$ fixed, X admits only finitely many jumps larger than a , so we can proceed by induction on the number N of such jumps. With no loss of generality we can arrange the jump times in increasing order, $T_1 < T_2 < \dots < T_N$. In case $N = 0$, $X = X^a$ and the claim holds. Suppose that

$$X^{n-1} = X^a + \sum_{i \leq n-1} \Delta X_{T_i} \mathbf{1}_{\{|\Delta X_{T_i}| > a\}} \mathbf{1}_{[[T_i, \infty[[$$

satisfies (29). Given that $X^n = X^{n-1}$ on $[[0, T_n[[$ and $X^c = X^{n,c} = X^{n-1,c}$, then if $t \leq T_n$ we have

$$\begin{aligned} I_{X^n}(h)_t &= I_{X^n}(h)_t - I_{X^n}(h)_{t-1/k} + I_{X^{n-1}}(h)_{t-1/k} \\ &= I_{X^n}(h)_t - I_{X^n}(h)_{t-1/k} + \int_0^{t-1/k} T_{X^{n-1}}^\mathcal{P}(h)d\langle X^c, X^c \rangle + \sum_{s \leq t-1/k} \Delta I_{X^{n-1}}(h)_s \\ &= \int_0^t T_{X^n}^\mathcal{P}(h)d\langle X^c, X^c \rangle + \sum_{s \leq t} \Delta I_{X^n}(h)_s \end{aligned}$$

the last equality being obtained by passing to the limit. On the other hand, let $T_n^p = T_n$ on $\{|\Delta X_{T_n}| < p\}$ or else $T_n^p = \infty$ and define $Y_t^p = X_{t \vee T_n^p}^n$ with $Y_\infty^p = 0$. Then, Y^p is locally bounded and therefore satisfies (29). Moreover, $X^n = Y^p$ on $[[T_n^p, \infty[[$. But then, if $t > T_n^p$ we have

$$\begin{aligned} I_{X^n}(h)_t &= I_{X^n}(h)_{T_n^p} + I_{Y^p}(h)_t \\ &= I_{X^n}(h)_{T_n^p} + \int_0^t T_{Y^p}^\mathcal{P}(h)d\langle Y^{p,c}, Y^{p,c} \rangle + \sum_{s \leq t} \Delta I_{Y^p}(h)_s \\ &= I_{X^n}(h)_{T_n^p} + \int_{T_n^p}^t T_{X^n}^\mathcal{P}(h)d\langle X^c, X^c \rangle + \sum_{T_n^p < s \leq t} \Delta I_{X^n}(h)_s \\ &= \int_0^t T_{X^n}^\mathcal{P}(h)d\langle X^c, X^c \rangle + \sum_{s \leq t} \Delta I_{X^n}(h)_s. \end{aligned}$$

The claim follows upon noting that $T_n^p \downarrow T_n$, a.s.. □

A one sided version of the condition (27) is also quite useful in applications.

Theorem 3. *Let $h \in \mathcal{H}_X$ be such that $I_X(h)$ is adapted and locally integrable and that there exists a sequence $\langle \sigma_n \rangle_{n \in \mathbb{N}}$ in \mathcal{T} which increases to ∞ a.s. and satisfies*

$$(34) \quad \lim_{c \rightarrow \infty} P^* \left(\inf_{s < t \leq \sigma_n} T_X(h) < -c \right) = 0 \quad n \in \mathbb{N}.$$

Then, $I_X(h)$ is a submartingale of locally integrable variation and admits a continuous process C^h of locally integrable variation such that

$$(35) \quad C^h \geq \int T_X^{\mathcal{P}}(h) d\langle X^c, X^c \rangle \quad \text{and} \quad I_X(h)_t = C_t^h + \sum_{s \leq t} \Delta I_X(h)_s.$$

If $T_X(h) \geq 0$ then C^h is increasing.

Proof. The proof is essentially the same as that of Theorem 4, so we only sketch the salient points. Restricting attention to a sequence $\langle A_{n,k} \rangle_{k \in \mathbb{N}}$ of \mathcal{F} measurable subsets of $\{\sup_{s < t \leq \sigma_n} T_X(h) \geq -k\}$ of arbitrary large probability, we conclude from $\psi(T_X(h)_{\tau \wedge \sigma_n} \mathbb{1}_{A_{n,k}}) \geq 0$ that $P(I_X(h)_\tau) \geq P \int_0^\tau T_X^{\mathcal{P}}(h) d\langle X^c, X^c \rangle$. The proof that $I_X(h)$ is of locally integrable variation follows from the inequality

$$P(|I_X(h)_\tau - I_X(h)_\sigma|) \leq \psi(T_X(h) \mathbb{1}_{[\sigma, \tau] \times \mathbb{R}_+}) + \int |T_X(h)| \mathbb{1}_{[\sigma, \tau] \times \mathbb{R}_+} d\mu.$$

We conclude that $I_X(h) - \int T_X^{\mathcal{P}}(h) d\langle X, X \rangle$ is a submartingale of locally integrable variation so its martingale part is just a compensated sum of jumps. We write it as $M + A$, with A an increasing, predictable process. Looking at its jumps we find that

$$\sum_{s \leq t} \Delta I_X(h)_s - T_X^{\mathcal{P}}(h)_s \Delta X_s^2 = M + \sum_{s \leq t} \Delta A_s$$

This proves the claim upon defining

$$(36) \quad C_t^h = A_t^c + \int_0^t T_X^{\mathcal{P}}(h) d\langle X^c, X^c \rangle$$

with A^c the continuous part of A . It is clear that if $T_X(h) \geq 0$ then C^h is increasing. \square

3. APPLICATIONS

In this section we specialize our preceding results, proving an extension of Îto's and Tanaka's formulas. Our proof is clearly inspired by [5]. The typical element $h \in \mathcal{H}_X$ will be of the form $h = f * g$ with

$$(37) \quad (f * g)(x, y) = f(y) - f(x) - g(x)(y - x) \quad f, g \in \mathfrak{F}(\mathbb{R}), \quad x, y \in \mathbb{R}.$$

Incidentally we note that $f * g \geq 0$ if and only if f is convex and $D_- f \leq g \leq D_+ f$ while $|f * g| \leq \hat{f} * \hat{g}$ for some pair $\hat{f}, \hat{g} \in \mathfrak{F}(\mathbb{R})$ if and only if f is the difference of two convex functions, g admits right and left limits and $D_\pm f = g_\pm$.

This remark together with Theorem 3 deliver an immediate proof of the original claim of Tanaka. In fact, when f is convex and $D_-f \leq g \leq D_+f$ (34) translates into

$$(38) \quad f(X_t) = f(X_0) + \int_0^t (g \circ X)_- dX + C_t^f + \sum_{s \leq t} \{ \Delta(f \circ X)_s - (g \circ X)_s \Delta X_s \}$$

where C^f is continuous, increasing and $C^f \geq \int T_X^{\mathcal{P}}(f * g) d\langle X^c, X^c \rangle$.

Example 1. Let $g \in \mathfrak{F}(\mathbb{R})$ be Lipschitz of rank $2K$ and f one of its primitives. Then,

$$|f(y) - f(x) - g(x)(y - x)| = \left| \int_x^y [g(t) - g(x)] dt \right| \leq K(y - x)^2$$

so that $|T_X(f * g)| \leq KT_X(h_0)$ and (27) applies.

Theorem 4. Let $g \in \mathcal{C}(\mathbb{R})$ be locally Lipschitz and f its primitive. Then,

$$(39) \quad \begin{aligned} f(X_t) = f(X_0) &+ \int_0^t (g \circ X)_- dX + \int_0^t T_X^{\mathcal{P}}(f * g) d\langle X^c, X^c \rangle \\ &+ \sum_{s \leq t} \{ \Delta(f \circ X)_s - (g \circ X)_s \Delta X_s \} \end{aligned}$$

where the sum $\sum_{s \leq \cdot} \{ \Delta(f \circ X)_s - (g \circ X)_s \Delta X_s \}$ describes a process of finite variation.

Proof. Put $h = f * g$, so that $I_X(h)_t = f(X_t) - f(X_0) - \int_0^t (g \circ X)_- dX$.

Assume initially that X is bounded and that $[X, X]_\tau$ and $\int_0^\tau (g \circ X)_- dX$ are integrable for each $\tau \in \mathcal{T}$, i.e. $\mathcal{T}(h_0) = \mathcal{T}$. Given compactness we can further assume that g is Lipschitz. Then, as we saw in Example 1, $T_X(h)$ is bounded and we obtain (39) from (29). The extension to the case in which X is just locally bounded is immediate.

As in the proof of Theorem 2, the extension from the case of a locally bounded semimartingale to that of a general semimartingale is done by noting first that for fixed $a > 0$ the process $X_t^a = X_t - \sum_{s \leq t} \Delta X_s \mathbf{1}_{\{|\Delta X_s| > a\}}$ satisfies the claim and then showing inductively that (39) is preserved when introducing a single, unbounded jump of which X has only a limited number on each bounded interval. \square

The following is a further extension in which, however, we cannot establish a precise extension formula.

Corollary 1. If $f \in \mathcal{C}^1(\mathbb{R})$ then, $f(X)$ is a semimartingale.

Proof. Assume again that X is bounded and let g be the derivative of f . Locally Lipschitz functions on a compact set K form a collection which contains the constants, is closed with respect to multiplication and separates points. There is then a sequence $\langle g_n \rangle_{n \in \mathbb{N}}$ of locally Lipschitz functions which converges uniformly to g . Let $f_n(x) = f(a) + \int_a^x g_n(t) dt$ for some $a \in K$. Then, f_n converges to f uniformly. By Theorem 4, f_n and g_n satisfy (39), so that, taking limits in probability as $n \rightarrow \infty$, we conclude that

$$f(X_t) = f(X_0) + \int_0^t (g \circ X)_- dX + \lim_n V_{n,t} = f(X_0) + \int_0^t (g \circ X)_- dX + V_t$$

where $V_n = \int T_X^{\mathcal{P}}(f_n * g_n) d\langle X^c, X^c \rangle + \sum_s \{\Delta(f_n \circ X)_s - (g_n \circ X)_{s-} \Delta X_s\}$ and converges in probability uniformly over compact sets to some limit V , since the other terms do. Then necessarily necessarily V is a semimartingale and so is $f \circ X$. \square

Corollary 2. *Let $g = g^c + g^d \in \mathfrak{F}(\mathbb{R})$ with g^c continuous and g^d of finite variation and let f be a primitive of g . Then, $f \circ X$ is a semimartingale.*

Proof. Write $f = f^c + f^d$ with f^c a primitive of g^c and f^d of g^d . Given that g^d is of finite variation, it splits into the difference of two increasing functions. Thus f^d splits into the difference of two convex functions to which (29) applies as a consequence of Tanaka's theorem. We obtain the same conclusion for f^c as this is an element of $\mathcal{C}^1(\mathbb{R})$, by Theorem 4. \square

REFERENCES

- [1] K. P. S. Bhaskara Rao, M. Bhaskara Rao (1983), *Theory of Charges*, Academic Press, London.
- [2] G. Cassese (2017), *Conglomerability and the Representation of Linear Functionals*, mimeo.
- [3] C. Dellacherie, P. A. Meyer (1982), *Probability and Potential*, North-Holland, Amsterdam.
- [4] N. Dunford, J. T. Schwartz (1988), *Linear Operators. Part I*, Wiley and Sons, New York.
- [5] P. A. Meyer (1976), *Un Cours sur les Integrales Stochastiques*, Sem. Probab. X, 17, Lect. Notes Math. 511, Springer-Verlag, Berlin-Heidelberg.

UNIVERSITÀ MILANO BICOCCA

E-mail address: gianluca.cassese@unimib.it

Current address: Department of Economics, Statistics and Management Building U7, Room 2097, via Bicocca degli Arcimboldi 8, 20126 Milano - Italy